DynamiCrafter: Animating Open-domain Images with Video Diffusion Priors

Year: 2023 | Citations: 390 | Authors: Jinbo Xing, Menghan Xia, Yong Zhang, Haoxin Chen, Gongye Liu

Abstract

Animating a still image offers an engaging visual experience. Traditional image animation techniques mainly focus on animating natural scenes with stochastic dynamics (e.g. clouds and fluid) or domain-specific motions (e.g. human hair or body motions), and thus limits their applicability to more general visual content. To overcome this limitation, we explore the synthesis of dynamic content for open-domain images, converting them into animated videos. The key idea is to utilize the motion prior of text-to-video diffusion models by incorporating the image into the generative process as guidance. Given an image, we first project it into a text-aligned rich context representation space using a query transformer, which facilitates the video model to digest the image content in a compatible fashion. However, some visual details still struggle to be preserved in the resultant videos. To supplement with more precise image information, we further feed the full image to the diffusion model by concatenating it with the initial noises. Experimental results show that our proposed method can produce visually convincing and more logical&natural motions, as well as higher conformity to the input image. Comparative evaluation demonstrates the notable superiority of our approach over existing competitors.